

Research Article

The Effect of Nonnative Interactions on the Energy Landscapes of Frustrated Model Proteins

Mark T. Oakley,¹ David J. Wales,² and Roy L. Johnston¹

¹ School of Chemistry, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

² University Chemical Laboratories, Lensfield Road, Cambridge CB2 1EW, UK

Correspondence should be addressed to Mark T. Oakley, m.t.oakley@bham.ac.uk

Received 15 November 2011; Accepted 27 January 2012

Academic Editor: Jan Petter Hansen

Copyright © 2012 Mark T. Oakley et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The 46- and 69-residue BLN model proteins both exhibit frustrated folding to β -barrel structures. We study the effect of varying the strength of nonnative interactions on the corresponding energy landscapes by introducing a parameter λ , which scales the potential between the BLN ($\lambda = 1$) and Gō-like ($\lambda = 0$) limits. We study the effect of varying λ on the efficiency of global optimisation using basin-hopping and genetic algorithms. We also construct disconnectivity graphs for these proteins at selected values of λ . Both methods indicate that the potential energy surface is frustrated for the original BLN potential but rapidly becomes less frustrated as λ decreases. For values of $\lambda \leq 0.9$, the energy landscape is funnelled. The fastest mean first encounter time for the global minimum does not correspond to the Gō model: instead, we observe a minimum when the favourable nonnative interactions are still present to a small degree.

1. Introduction

Proteins are biopolymers constructed from a sequence of amino acid residues. The potential energy landscapes of proteins have many degrees of freedom and include important contributions between pairs of residues that are distant in sequence, but close to each other in space. Despite this complexity, many globular proteins fold to a well-defined native state. According to the thermodynamic hypothesis, this structure is the global free energy minimum for a given sequence [1]. Frustration occurs when there are low-lying structures separated by high barriers [2]. All the favourable interactions between pairs of residues cannot be accommodated at the same time, which can lead to energetic frustration, where there are several low-lying structures with different patterns of contacts. Geometric frustration occurs when the interconversion of two low-lying structures requires the breaking of several favourable contacts.

A systematic way to simplify the potential energy surface for a protein is to include only attractive interactions between pairs of residues that are in contact in the native state, which constitutes a Gō model [3]. Various on- and off-lattice Gō models have been investigated by different authors to

study a range of different proteins. In spite of the simplified potential, these models have proved capable of reproducing certain aspects of protein dynamics and thermodynamics [4–11]. Using a Gō model tends to lead to funnelled energy landscapes [12], with very little frustration. For some proteins, neglecting nonnative interactions can have a significant influence on the energy landscape [13].

United atom representations introduce a further level of coarse-graining, which can speed up simulations significantly, at the cost of atomistic detail. The simplest coarse-grained model is the HP model, in which each protein residue is represented by a single hydrophobic (H) or polar (P) bead and is constrained to lie on a regular lattice [14, 15]. The BLN model is an off-lattice generalisation of the HP model with three types of bead: hydrophobic (B), hydrophilic (L), and neutral (N). The 46-residue sequence [12, 16–33] $B_9N_3(LB)_4N_3B_9N_3(LB)_5L$ and the 69-residue sequence [34–38] $B_9N_3(LB)_4N_3B_9N_3(LB)_4N_3B_9N_3(LB)_5L$ were designed to exhibit frustrated folding and have several alternate β -barrel structures that are separated by large energy barriers. Disconnectivity graphs [39] for both of these proteins exhibit energy landscapes comprising several folding funnels [12, 38]. Using a Gō potential for these two proteins changes

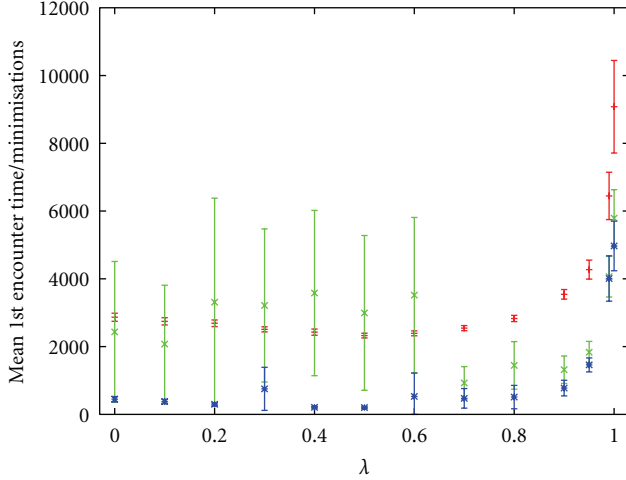


FIGURE 1: Mean first encounter times (number of minimisations) for 100 global optimisation runs initiated from random starting points for the 46-residue scaled BLN protein. The searches were run using a genetic algorithm (red), basin-hopping starting from random structures confined to a sphere (green), and basin-hopping starting from chain structures with randomised dihedral angles (blue). The error bars are the uncertainties calculated at the 95% level.

the nature of their energy landscapes, and they both exhibit single funnels with very little frustration [12, 38].

Intermediate potentials can be generated using a parameter, λ , which scales the strength of the nonnative interactions between the Gō ($\lambda = 0$) and BLN ($\lambda = 1$) limits. The folding thermodynamics of the 46-residue BLN protein have been investigated using this scaled BLN potential [23, 32, 33], showing that most of the frustration is only present for values of $\lambda \geq 0.9$. The introduction of salt bridges (gatekeepers) to the 46-residue protein also produces energy landscapes of intermediate character [27, 28].

In the present work, we study the effect of varying λ on the ease of global optimisation of the 46- and 69-residue BLN proteins using a basin-hopping algorithm and a genetic algorithm. We also construct disconnectivity graphs to compare the energy landscapes of the proteins for different values of λ .

2. Computational Methods

The protein structures were modelled using the following BLN potential [12, 21, 26, 28]:

$$\begin{aligned}
 V_{\text{BLN}} = & \frac{1}{2} K_r \sum_{i=1}^{N-1} (R_{i,i+1} - R_e)^2 + \frac{1}{2} K_\theta \sum_{i=1}^{N-2} (\theta_i - \theta_e)^2 \\
 & + \epsilon \sum_{i=1}^{N-3} [A_i(1 + \cos \phi_i) + B_i(1 + 3 \cos \phi_i)] \\
 & + 4\epsilon \sum_{i=1}^{N-2} \sum_{j=i+2}^N C_{ij} \left[\left(\frac{\sigma}{R_{ij}} \right)^{12} - D_{ij} \left(\frac{\sigma}{R_{ij}} \right)^6 \right],
 \end{aligned} \quad (1)$$

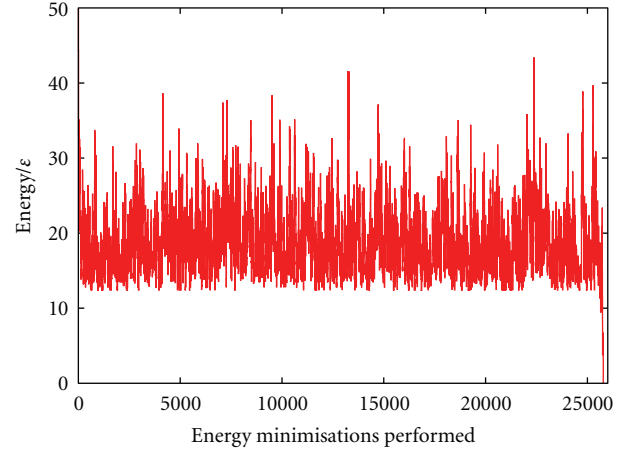


FIGURE 2: Energy of the minima in the Markov chain for a BH run where trapping occurs for the 46-residue scaled BLN protein with $\lambda = 0$.

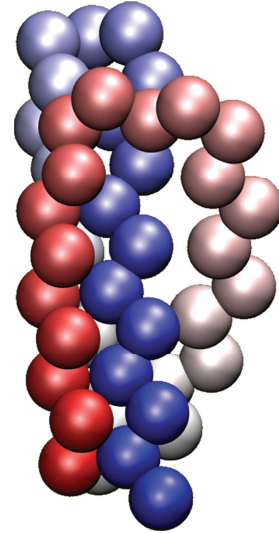


FIGURE 3: The most stable misfolded structure, which acts as a trap for global optimisation of the 46-residue BLN protein, illustrated using the VMD program [40] with a colouring scheme for the beads that varies from red to blue (N-terminus to C-terminus).

where R_{ij} is the distance between two beads i and j . The first term is a harmonic bond restraint with $K_r = 231.2\epsilon\sigma^{-2}$ and $R_e = \sigma$. The second term is a bond angle restraint with $K_\theta = 20 \text{ rad}^{-2}$ and $\theta_e = 1.8326 \text{ rad}$. The third term involves torsional angles, ϕ , defined by four successive beads. If two or more of these beads are N, then $A = 0$ and $B = 0.2$. For all other sequences, $A = B = 1.2$. The final term introduces pairwise nonbonded interactions. If one residue is L and the other is L or B, then $C = 2/3$ and $D = -1$. If either of the residues is N, then $C = 1$ and $D = 0$. If both residues are B, then $C = 1$, but the value of D depends on the presence of the contact in the native state of the protein. For native contacts, $D = 1$. For nonnative contacts, $D = \lambda$, where $0 < \lambda < 1$. The case where $\lambda = 1$ is the original BLN potential and $\lambda = 0$ is the Gō potential.

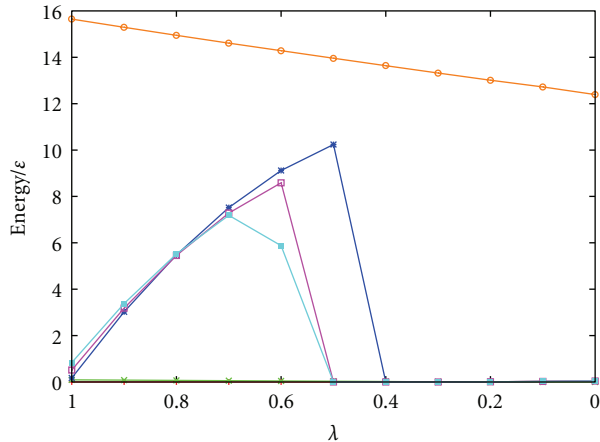


FIGURE 4: The energies of the five most stable BLN-46 structures relative to the global minimum as a function of λ . Also shown (orange) is the energy of the trap structure illustrated in Figure 3. The steep decreases mark the points at which structures cease to be local minima and collapse into the basin of attraction [41] of the global minimum.

TABLE 1: Parameters used for the two optimisation strategies.

	BLN-46	BLN-69
BH		
kT/ϵ	2.3	3.4
Step size/ σ	0.65	0.70
GA		
Population size	140	200
Crossover rate	0.9	0.9
Mutation rate	0.05	0.05

Native contacts are defined as all pairs of residues where R_{ij} is less than a fixed cut-off distance in the native state (global minimum) of the protein. When $\lambda \neq 1$, the value of this cut-off radius will influence the energy landscape. Here, we use 1.167σ for consistency with previous work [12, 28, 38].

Global optimisation was performed using the basin-hopping approach [42–44] and a Lamarckian genetic algorithm [38, 45], which are both implemented in the GMIN program [46]. Each algorithm involves local energy minimisation after each structural perturbation. This minimisation transforms the potential energy surface into the basins of attraction of local minima [47] and removes downhill barriers. The search parameters for both algorithms were optimised in previous work for BLN proteins [38], and these parameters were used without adjustment for all searches presented here (Table 1). The GMIN input files used for these searches are included in the supplementary data (see Supplementary Material available online at doi:10.1155/2012/192613).

The genetic algorithm represents each structure with a genome consisting of the torsion angles in the backbone of the protein. Offspring structures are generated by one-point crossover from two parent structures. Mutants are

generated by making a copy of an existing structure (parent or offspring) and replacing one of the torsion angles. To prevent stagnation of the genetic algorithm searches, a restart operator was used. If an entire generation of offspring contains no solutions that are fitter than any of the parent structures, a new epoch is started with a new random population. For the 69-residue protein, the fittest structure from each epoch survives into the next epoch.

All conformational searches were run until the global minimum structure was found. We report the mean time taken to encounter this structure in conformational searches from randomised starting points to compare the exploration of the energy landscape as a function of λ . Searches were performed for values of λ between 0 and 1 in steps of 0.1, with additional points at $\lambda = 0.95$ and $\lambda = 0.99$. The initial structures for this benchmarking were generated using two alternative methods: either random placement of the residues inside a sphere of radius 3σ , or random assignment of the backbone dihedral angles. Full details of all of the global optimisation runs are available as supplementary data.

The disconnectivity graphs for the model proteins were constructed from databases of stationary points generated using the PATHSAMPLE program, [48] which organises independent pathway searches using OPTIM [49]. All the transition state searches in OPTIM were conducted in Cartesian coordinates [50] using a quasicontinuous interpolation scheme to avoid chain crossings, with local maxima accurately refined to transition states by hybrid eigenvector-following [51–53]. Successive pairs of local minima were selected for connection attempts within OPTIM using the missing connection algorithm [54]. Disconnectivity graphs [39] will be illustrated for both the 46- and 69-residue scaled BLN proteins with λ values of 0, 0.5, 0.9, and 1.

We also study the effect of λ on key structures of the BLN proteins. These structures were reminimised using values of λ between 0 and 1 in steps of 0.1. Pathways between pairs of interesting minima were studied by Dijkstra analysis [55] in PATHSAMPLE [48], with the discrete paths [56] that make the largest contribution to the steady-state rate constant [56, 57] presented here.

With a few exceptions, all of the stationary points of the BLN model proteins are chiral. However, the BLN potential includes no chiral terms, so each structure has an enantiomer with the same energy. When evaluating the optimisation algorithms, we accept convergence to either of the enantiomers of the global minimum. When looking at the pathways, it is important to use the same chirality for both structures, otherwise much longer paths result. For some of the trapped structures, pathways to both enantiomers of the global minimum can be viable.

3. Results

3.1. BLN-46. Searches for $\lambda = 0$ (Gō potential) find the global minimum much more rapidly than when $\lambda = 1$ (BLN potential), as one would expect for a more funnelled energy landscape [2, 58–61]. However, the number of steps required varies nonlinearly between these two extremes and behaves

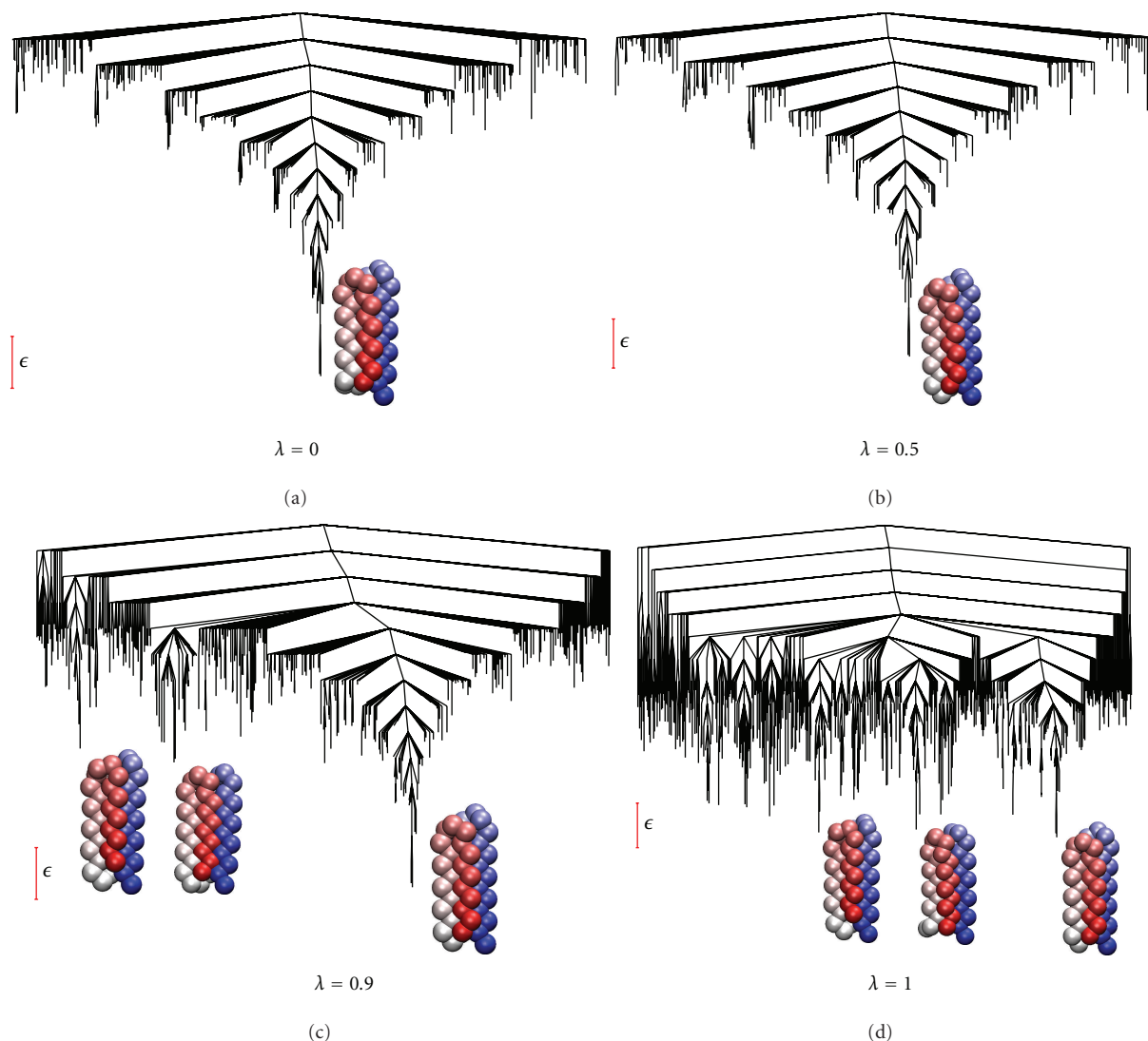


FIGURE 5: Disconnectivity graphs [39] showing the most stable minima accessible by transition states lower than 7ϵ from the global minimum of the 46-residue scaled BLN proteins. For $\lambda \geq 0.9$, only the 1000 most stable minima are shown. The structures of selected minima are illustrated close to the bottoms of the corresponding branches.

differently for each search algorithm. When optimising with the GA, the mean first encounter time decreases rapidly from $\lambda = 1$ to $\lambda = 0.9$ and then more slowly to a minimum at $\lambda = 0.5$ (Figure 1). After this minimum, there is a small increase in the required time as λ decreases to 0. This result is consistent with previous observations that the introduction of some nonnative interactions can assist the folding of some proteins [62]. Below $\lambda = 0.9$, almost all searches find the global minimum within the first epoch of the GA. For larger values of λ , several searches require two or more epochs, leading to much more variation in the first encounter time. The choice of the random starting configurations for the initial population of the GA makes little difference to the mean first encounter time.

In basin-hopping searches, the choice of starting structures makes a large difference to the efficiency of the optimisation. When starting from residues randomly distributed inside a sphere, for values of $\lambda < 0.7$, 95% of the searches

find the global minimum rapidly. The remaining searches become trapped and require several thousand attempted Monte Carlo moves to escape (Figure 2). In this trap, the first, third, and fourth strands are correctly packed, but the second is wrapped around the outside of the protein (Figure 3). Searches with larger values of λ do not become trapped in this basin, which suggests that the nonnative interactions are important in stabilising the intermediates between this structure and the global minimum.

The trap configuration lies 12.4ϵ above the global minimum when $\lambda = 0$ and becomes more unfavourable for larger values of λ (Table 2). The fastest escape route from this trap involves unthreading of the N-terminus from the loop made by the second strand (Table 2). The energy of the highest transition state on this pathway relative to the trapped state increases from $\lambda = 0$ to $\lambda = 0.9$ before levelling off. The highest transition state on this pathway lies above the barrier to interconversion of the two enantiomers of

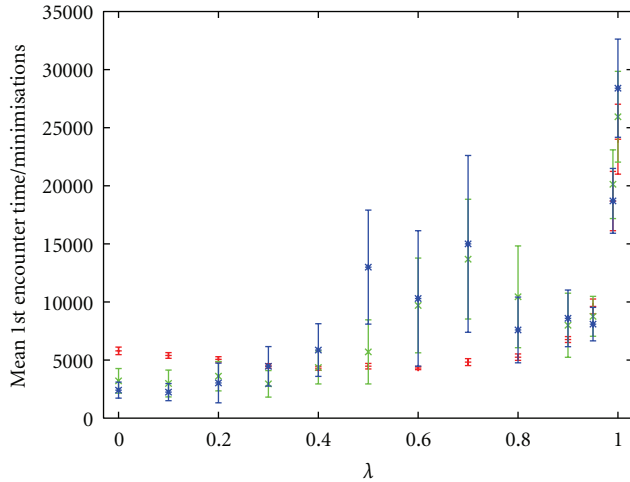


FIGURE 6: Mean first encounter times (number of minimisations) for 100 global optimisation runs initiated from random starting points for the 69-residue scaled BLN protein. The searches were run using a genetic algorithm (red), basin-hopping starting from random structures confined to a sphere (green), and basin-hopping starting from chain structures with randomised dihedral angles (blue). The error bars are the uncertainties calculated at the 95% level.

the global minimum. For searches starting from a random set of torsion angles, this trapping is much less frequent and is only seen in 3 of the 700 searches performed where $0 \leq \lambda \leq 0.6$. By retaining some notion of connectivity, these initial structures cover less of the configurational space than the entirely random starting points. However, the complete coverage of conformational space comes at the cost of including more unstable structures, such as the trap seen here.

The five lowest minima in the BLN-46 protein span an energy range of less than ϵ (Figure 4). The two most stable minima are in the same basin, and both have all of the BB contacts from the native state. Across the range of λ , the relative energies of these minima are within 0.1ϵ of each other, with the second-best minimum becoming slightly more stable as λ decreases and moving below the former global minimum when $\lambda < 0.3$ [12]. The next three minima are stabilised by some nonnative contacts and become less stable relative to the global minimum as λ decreases. In the region around $\lambda = 0.5$, these structures cease to be minima and fall into the basins of attraction [41] of the two lowest energy structures.

The disconnectivity graphs within 7ϵ of the global minimum for $\lambda = 0$ and $\lambda = 0.5$ are funnelled and almost indistinguishable (Figure 5). When $\lambda = 0.9$, some frustration appears in the low-energy regions of the energy landscape, but it is still mostly funnelled. Almost all of the frustration is introduced between $\lambda = 0.9$ and $\lambda = 1$, where several alternate β -barrel structures are separated by barriers of 4 to 5ϵ . This organisation is consistent with the increase in the mean first encounter times seen for global optimisation with $\lambda > 0.9$ and agrees with previous studies of the thermodynamics of the 46-residue protein [32, 33],

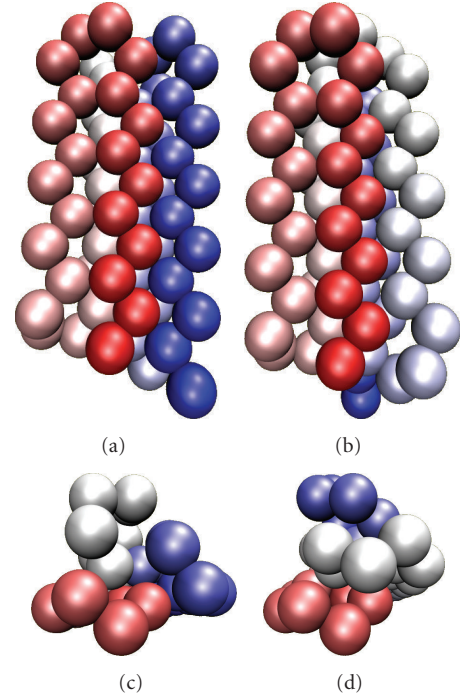


FIGURE 7: Side and top views of the global minimum (left) and trapped (right) structures of the 69-residue BLN protein illustrated using the VMD program [40] with a colouring scheme for the beads that varies from red to blue (N-terminus to C-terminus).

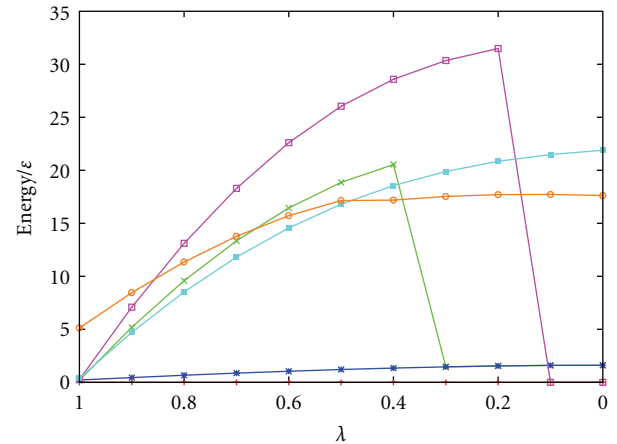


FIGURE 8: The energies of the five most stable BLN-69 structures relative to the global minimum as a function of λ . Also shown (orange) is the trap structure from Figure 7. The steep decreases in energy mark the points at which structures cease to be local minima and collapse into the basin of attraction [41] of the global minimum.

where $\lambda = 0$ and $\lambda = 0.5$ were found to be good folders, $\lambda = 0.9$ an intermediate folder, and $\lambda = 1$ a poor folder.

3.2. BLN-69. The behaviour of the GA for the 69-residue protein is similar to that for the 46-residue protein, with the fastest search time found at $\lambda = 0.5$. When optimising with basin-hopping on the 69-residue protein, there are several

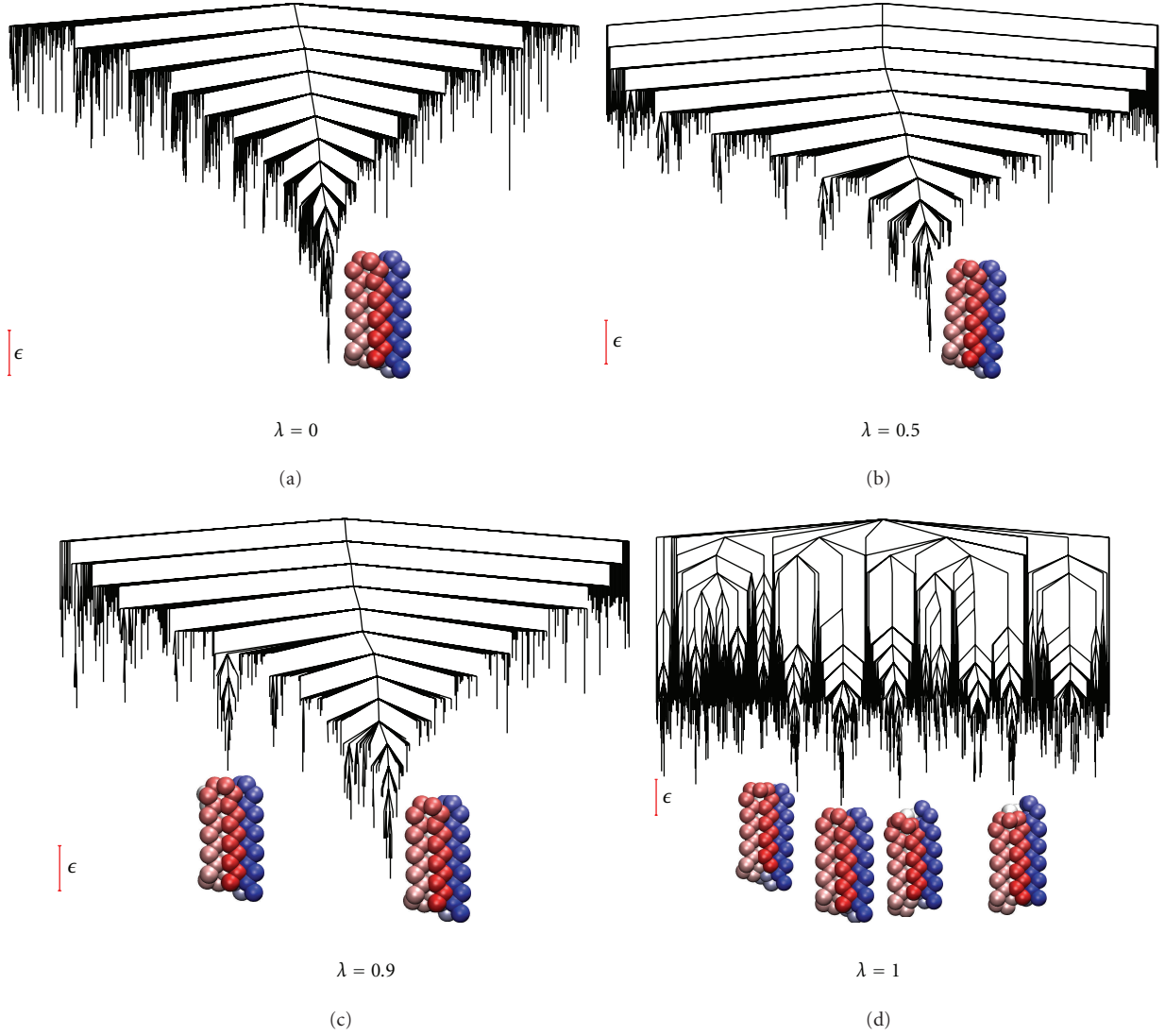


FIGURE 9: Disconnectivity graphs [39] showing the minima accessible by transition states lower than 8ϵ from the global minimum of the 69-residue scaled BLN proteins. For $\lambda \geq 0.5$, only the 1000 most stable minima are shown. The structures of selected minima are illustrated close to the bottoms of the corresponding branches.

TABLE 2: Energies of the trapped minimum and transition state for escape from the principal kinetic trap in the 46-residue scaled BLN protein. All energies are in units of ϵ and measured relative to the global minimum.

λ	E_{trap}	E_{untrap}
0.0	12.4	22.9
0.5	14.0	27.1
0.9	15.3	32.2
1.0	15.6	32.2

slow searches between $\lambda = 0.4$ and $\lambda = 0.8$ (Figure 6). There are multiple trap structures, and the one that is seen most frequently, which is responsible for the slowest searches, is formed from three strands from the left-handed barrel and three strands from right-handed barrel (Figure 7). This

structure is a six-stranded β -barrel similar to the global minimum, but with two sets of interstrand contacts swapped (1–6 and 3–4 in the global minimum compared to 1–4 and 3–6 in the trap).

Conversion from the above structure to the global minimum proceeds either by inversion of the three strands at the N-terminus or of the three strands at the C-terminus. The barriers to these two mechanisms are different and vary with λ (Table 3). The barrier for the fastest pathway for inversion at the C-terminus becomes larger with increasing λ . However, the barrier for inversion of the N-terminus varies much less with λ . In the region where $0.5 \leq \lambda \leq 0.7$, the barriers to both routes out of the trap are relatively high, which is a possible explanation for the slow basin-hopping optimisation for these values of λ . This is doubtless an oversimplification when we consider that there are multiple trap structures.

TABLE 3: Energies of the trapped minimum and transition states for escape from the principal kinetic trap by inversion of the N- and C-termini in the 69-residue scaled BLN protein. All energies are in units of ϵ and measured relative to the global minimum.

λ	E_{trap}	$E_{\text{untrap-C}}$	$E_{\text{untrap-N}}$
0.0	17.6	30.2	38.2
0.5	16.7	43.8	33.2
0.7	13.8	30.2	31.1
0.9	8.4	25.7	26.9
1.0	5.1	25.8	23.9

For the 69-residue BLN protein, the energies of the five lowest minima span less than 0.4ϵ (Figure 8). One structure lies in the same funnel as the global minimum, and its relative energy increases from 0.2ϵ to 1.6ϵ when λ decreases from 1 to 0. The other three structures occupy different funnels from the global minimum, with several nonnative contacts, and their stability decreases steeply with decreasing λ . Unlike the 46-residue protein, the global minimum structure remains the same for all values of λ . The low-energy region of disconnectivity graphs for values of λ between 0 and 0.9 are mostly funnelled (Figure 9). Almost all of the frustration in this region of the potential energy surface appears for $\lambda > 0.9$.

4. Conclusions

Much of the energetic frustration in the BLN proteins is removed once the potential contains a 10% contribution from the Gō function. When looking at geometric frustration in higher-energy traps, the effect of λ is less predictable. The removal of nonnative interactions can stabilise or destabilise the transition states that must be crossed to escape from these traps. Measures of the landscape complexity [30] could provide a useful way to understand the influence of nonnative interactions and will be considered in future work.

Acknowledgments

The authors acknowledge the Engineering and Physical Sciences Research Council, UK (EPSRC), for funding under Programme Grant EP/I001352/1. The calculations described in this paper were performed using the University of Birmingham's BlueBEAR HPC service, which was purchased through HEFCE SRIF-3 funds (see <http://www.bear.bham.ac.uk/>).

References

- [1] C. B. Anfinsen, "Principles that govern the folding of protein chains," *Science*, vol. 181, no. 4096, pp. 223–230, 1973.
- [2] J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, "Funnel, pathways, and the energy landscape of protein folding: a synthesis," *Proteins*, vol. 21, no. 3, pp. 167–195, 1995.
- [3] Y. Ueda, H. Taketomi, and N. Gō, "Studies on protein folding, unfolding, and fluctuations by computer simulation. II. A three-dimensional lattice model of lysozyme," *Biopolymers*, vol. 17, no. 6, pp. 1531–1548, 1978.
- [4] C. Micheletti, F. Seno, and A. Maritan, "Polymer principles of protein calorimetric two-state cooperativity," *Proteins*, vol. 40, no. 4, pp. 637–661, 2000.
- [5] C. Clementi, H. Nymeyer, and J. N. Onuchic, "Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins," *Journal of Molecular Biology*, vol. 298, no. 5, pp. 937–953, 2000.
- [6] J. W. H. Schymkowitz, F. Rousseau, and L. Serrano, "Surfing on protein folding energy landscapes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 25, pp. 15846–15848, 2002.
- [7] P. Das, C. J. Wilson, G. Fossati, P. Wittung-Stafshede, K. S. Matthews, and C. Clementi, "Characterization of the folding landscape of monomeric lactose repressor: quantitative comparison of theory and experiment," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 41, pp. 14569–14574, 2005.
- [8] A. R. Lam, J. M. Borreguero, F. Ding et al., "Parallel folding pathways in the SH3 domain protein," *Journal of Molecular Biology*, vol. 373, no. 5, pp. 1348–1360, 2007.
- [9] P. F. N. Faísca, R. D. M. Travasso, R. C. Ball, and E. I. Shakhnovich, "Identifying critical residues in protein folding: insights from ϕ -value and P_{fold} analysis," *Journal of Chemical Physics*, vol. 129, no. 9, Article ID 095108, 2008.
- [10] R. D. Hills and C. L. Brooks, "Insights from coarse-grained go models for protein folding and dynamics," *International Journal of Molecular Sciences*, vol. 10, no. 3, pp. 889–905, 2009.
- [11] P. O. Craig, J. Lätzer, P. Weinkam et al., "Prediction of native-state hydrogen exchange from perfectly funneled energy landscapes," *American Chemical Society*, vol. 133, no. 43, pp. 17463–17472, 2011.
- [12] M. A. Miller and D. J. Wales, "Energy landscape of a model protein," *Journal of Chemical Physics*, vol. 111, no. 14, pp. 6610–6616, 1999.
- [13] L. Sutto, J. Lätzer, J. A. Hegler, D. U. Ferreira, and P. G. Wolynes, "Consequences of localized frustration for the folding mechanism of the IM7 protein," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 50, pp. 19825–19830, 2007.
- [14] K. F. Lau and K. A. Dill, "A lattice statistical mechanics model of the conformational and sequence spaces of proteins," *Macromolecules*, vol. 22, no. 10, pp. 3986–3997, 1989.
- [15] K. A. Dill, S. Bromberg, K. Yue et al., "Principles of protein folding—a perspective from simple exact models," *Protein Science*, vol. 4, no. 4, pp. 561–602, 1995.
- [16] J. D. Honeycutt and D. Thirumalai, "Metastability of the folded states of globular proteins," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 87, no. 9, pp. 3526–3529, 1990.
- [17] J. D. Honeycutt and D. Thirumalai, "The nature of folded states of globular proteins," *Biopolymers*, vol. 32, no. 6, pp. 695–709, 1992.
- [18] Z. Guo and D. Thirumalai, "Nucleation mechanism for protein folding and theoretical predictions for hydrogen-exchange labeling experiments," *Biopolymers*, vol. 35, no. 1, pp. 137–140, 1995.
- [19] Z. Guo and D. Thirumalai, "Kinetics and thermodynamics of folding of a de novo designed four-helix bundle protein," *Journal of Molecular Biology*, vol. 263, no. 2, pp. 323–343, 1996.
- [20] Z. Guo and C. L. Brooks III, "Thermodynamics of protein folding: a statistical mechanical study of a small all- β protein," *Biopolymers*, vol. 42, no. 7, pp. 745–757, 1997.

- [21] R. S. Berry, N. Elmáci, J. P. Rose, and B. Vekhter, "Linking topography of its potential surface with the dynamics of folding of a protein model," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 18, pp. 9520–9524, 1997.
- [22] H. Nymeyer, A. E. García, and J. N. Onuchic, "Folding funnels and frustration in off-lattice minimalist protein landscapes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 11, pp. 5921–5928, 1998.
- [23] J. E. Shea, Y. D. Nochomovitz, Z. Guo, and C. L. Brooks, "Exploring the space of protein folding Hamiltonians: the balance of forces in a minimalist β -barrel model," *Journal of Chemical Physics*, vol. 109, no. 7, pp. 2895–2903, 1998.
- [24] N. Elmáci and R. S. Berry, "Principal coordinate analysis on a protein model," *Journal of Chemical Physics*, vol. 110, no. 21, pp. 10606–10622, 1999.
- [25] J. E. Shea, J. N. Onuchic, and C. L. Brooks, "Energetic frustration and the nature of the transition state in protein folding," *Journal of Chemical Physics*, vol. 113, no. 17, pp. 7663–7671, 2000.
- [26] D. A. Evans and D. J. Wales, "Free energy landscapes of model peptides and proteins," *Journal of Chemical Physics*, vol. 118, no. 8, pp. 3891–3897, 2003.
- [27] A. D. Stoycheva, J. N. Onuchic, and C. L. Brooks, "Effect of gatekeepers on the early folding kinetics of a model β -barrel protein," *Journal of Chemical Physics*, vol. 119, no. 11, pp. 5722–5729, 2003.
- [28] D. J. Wales and P. E. J. Dewsbury, "Effect of salt bridges on the energy landscape of a model protein," *Journal of Chemical Physics*, vol. 121, no. 20, pp. 10284–10290, 2004.
- [29] T. Komatsuzaki, K. Hoshino, Y. Matsunaga, G. J. Rylance, R. L. Johnston, and D. J. Wales, "How many dimensions are required to approximate the potential energy landscape of a model protein?" *Journal of Chemical Physics*, vol. 122, no. 8, Article ID 084714, pp. 1–9, 2005.
- [30] G. J. Rylance, R. L. Johnston, Y. Matsunaga, C.-B. Li, A. Baba, and T. Komatsuzaki, "Topographical complexity of multidimensional energy landscapes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 49, pp. 18551–18555, 2006.
- [31] J. Kim and T. Keyes, "Inherent structure analysis of protein folding," *Journal of Physical Chemistry B*, vol. 111, no. 10, pp. 2647–2657, 2007.
- [32] J. Kim and T. Keyes, "Influence of go-like interactions on global shapes of energy landscapes in β -barrel forming model proteins: inherent structure analysis and statistical temperature molecular dynamics simulation," *Journal of Physical Chemistry B*, vol. 112, no. 3, pp. 954–966, 2008.
- [33] J. Kim, T. Keyes, and J. E. Straub, "Relationship between protein folding thermodynamics and the energy landscape," *Physical Review E*, vol. 79, no. 3, Article ID 030902, 2009.
- [34] S. A. Larrass, L. M. Pegram, H. L. Gordon, and S. M. Rothstein, "Efficient generation of low-energy folded states of a model protein. II. Automated histogram filtering," *Journal of Chemical Physics*, vol. 119, no. 24, pp. 13149–13158, 2003.
- [35] P. W. Pan, H. L. Gordon, and S. M. Rothstein, "Local-structural diversity and protein folding: application to all-beta off-lattice protein models," *The Journal of Chemical Physics*, vol. 124, no. 2, p. 024905, 2006.
- [36] J. Kim, J. E. Straub, and T. Keyes, "Statistical temperature molecular dynamics: application to coarse-grained β -barrel-forming protein models," *Journal of Chemical Physics*, vol. 126, no. 13, Article ID 135101, 2007.
- [37] S.-Y. Kim, "An off-lattice frustrated model protein with a six-stranded β -barrel structure," *Journal of Chemical Physics*, vol. 133, no. 13, Article ID 135102, 2010.
- [38] M. T. Oakley, D. J. Wales, and R. L. Johnston, "Energy landscape and global optimization for a frustrated model protein," *Journal of Physical Chemistry B*, vol. 115, no. 39, pp. 11525–11529, 2011.
- [39] O. M. Becker and M. Karplus, "The topology of multidimensional potential energy surfaces: theory and application to peptide structure and kinetics," *Journal of Chemical Physics*, vol. 106, no. 4, pp. 1495–1517, 1997.
- [40] W. Humphrey, A. Dalke, and K. Schulten, "VMD: visual molecular dynamics," *Journal of Molecular Graphics*, vol. 14, no. 1, pp. 33–38, 1996.
- [41] P. G. Mezey, *Potential Energy Hypersurfaces*, Elsevier, Amsterdam, The Netherlands, 1987.
- [42] Z. Li and H. A. Scheraga, "Monte Carlo-minimization approach to the multiple-minima problem in protein folding," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 84, no. 19, pp. 6611–6615, 1987.
- [43] D. J. Wales and H. A. Scheraga, "Global optimization of clusters, crystals, and biomolecules," *Science*, vol. 285, no. 5432, pp. 1368–1372, 1999.
- [44] D. J. Wales and J. P. K. Doye, "Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms," *Journal of Physical Chemistry A*, vol. 101, no. 28, pp. 5111–5116, 1997.
- [45] R. L. Johnston, "Evolving better nanoparticles: genetic algorithms for optimising cluster geometries," *Dalton Transactions*, no. 22, pp. 4193–4207, 2003.
- [46] D. J. Wales, "GMN: A program for finding global minima and calculating thermodynamic properties from basin-sampling," <http://www-wales.ch.cam.ac.uk/GMIN/>.
- [47] P. G. Mezey, "Catchment region partitioning of energy hypersurfaces, I," *Theoretica Chimica Acta*, vol. 58, no. 4, pp. 309–330, 1981.
- [48] D. J. Wales, "PATHSAMPLE: A program for refining and analysing kinetic transition networks," <http://www-wales.ch.cam.ac.uk/OPTIM/>.
- [49] D. J. Wales, "OPTIM: A program for characterising stationary points and reaction pathways," <http://www-wales.ch.cam.ac.uk/PATHSAMPLE/>.
- [50] D. J. Wales, "Locating stationary points for clusters in cartesian coordinates," *Journal of the Chemical Society, Faraday Transactions*, vol. 89, no. 9, pp. 1305–1313, 1993.
- [51] L. J. Munro and D. J. Wales, "Defect migration in crystalline silicon," *Physical Review B*, vol. 59, no. 6, pp. 3969–3980, 1999.
- [52] G. Henkelman and H. Jónsson, "A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives," *Journal of Chemical Physics*, vol. 111, no. 15, pp. 7010–7022, 1999.
- [53] Y. Kumeda, L. J. Munro, and D. J. Wales, "Transition states and rearrangement mechanisms from hybrid eigenvector-following and density functional theory: application to $C_{10}H_{10}$ and defect migration in crystalline silicon," *Chemical Physics Letters*, vol. 341, no. 1–2, pp. 185–194, 2001.
- [54] J. M. Carr, S. A. Trygubenko, and D. J. Wales, "Finding pathways between distant local minima," *Journal of Chemical Physics*, vol. 122, no. 23, Article ID 234903, pp. 1–7, 2005.
- [55] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, no. 1, pp. 269–271, 1959.
- [56] D. J. Wales, "Discrete path sampling," *Molecular Physics*, vol. 100, no. 20, pp. 3285–3305, 2002.

- [57] D. J. Wales, “Energy landscapes: calculating pathways and rates,” *International Reviews in Physical Chemistry*, vol. 25, no. 1-2, pp. 237–282, 2006.
- [58] J. D. Bryngelson and P. G. Wolynes, “Spin glasses and the statistical mechanics of protein folding,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 84, no. 21, pp. 7524–7528, 1987.
- [59] J. N. Onuchic, P. G. Wolynes, Z. Luthey-Schulten, and N. D. Socci, “Toward an outline of the topography of a realistic protein-folding funnel,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no. 8, pp. 3626–3630, 1995.
- [60] M. Karplus and A. Sali, “Theoretical studies of protein folding and unfolding,” *Current Opinion in Structural Biology*, vol. 5, no. 1, pp. 58–73, 1995.
- [61] J. N. Onuchic, H. Nymeyer, A. E. García, J. Chahine, and N. D. Socci, “The energy landscape theory of protein folding: insights into folding mechanisms and scenarios,” *Advances in Protein Chemistry*, vol. 53, pp. 87–152, 2000.
- [62] C. Clementi and S. S. Plotkin, “The effects of nonnative interactions on protein folding rates: theory and simulation,” *Protein Science*, vol. 13, no. 7, pp. 1750–1766, 2004.